

The Interaction of Linguistic and Social Factors in Orthography Development: The Case of Anii

Deborah Morton, The Ohio State University

morton@ling.ohio-state.edu

Stefanie Zaske, SIL Togo-Benin

stefanie_zaske@sil.org

ICLDC 3 Presentation

February 28, 2013

1.0 Introduction

Anii is a West African Kwa language spoken along the border between Togo and Benin by approximately 50,000 people. There are four major dialects which are quite different from each other in both pronunciation and grammar.

In 2012, an orthography for Anii was finalized through collaboration among linguists, literacy workers, and the Anii community. This presentation will discuss some of the linguistic and social factors that played a role in the orthographical decisions that were made.

1.1 Background on Literacy in Anii

Since the 1970s, the government of Benin, where most of the Anii villages are located, has been very supportive of the idea of literacy in local languages, going as far as paying literacy teachers in each village, and even sponsoring literacy classes. An alphabet was developed, but no orthography, and classes generally didn't go beyond teaching the alphabet, especially as there was very little available to read, and no standardized spelling. However, some Anii people remained very invested in the development of their language. There was even an attempt to start a magazine in the 1980s, though only two issues came out. All this is to say that the Anii people are proud of their language and interested in writing it, though the practicalities of doing so were always a barrier.

Since 2001 when an SIL project was established for the Anii language, efforts to make literature in Anii available to readers have been underway, and a big part of that has been the development of the orthography, including linguistic research and getting the community involved in orthography decisions.

1.2 Who Made Orthography Decisions

The orthography decisions which will be discussed below were taken largely at an orthography meeting in March of 2012 which was organized by the SIL team. Some follow-up decisions were taken by the members of the research group created at that meeting to finalize the remaining decisions.

Those who were present at the meeting were representatives from the following groups:

- The old men who have been involved with language development in Anii since the beginning
- The literacy teachers from many of the villages, representing all the major dialect groups
- The language committee, a group of local leaders who were chosen by the old men to continue the work of language development (including Anii-language radio personalities)
- The group of young people who are being trained to write and publish *GɔGɔ*, the Anii-language magazine (www.revue-gugu.org) that SIL has recently started publishing
- Other interested people, including government officials, people involved in development for the area, and the SIL team

With this background in mind, I will now present some of the decisions that were made.

2.0 Linguistic Factors

Some orthography decisions were based largely on the linguistic characteristics of the language. Examples of these types of decisions are given in this section.

2.1 Word boundaries

There are two linguistic arguments as to what constitutes a word in Anii. The first is, obviously, semantic, which suggest that each noun, each verb, each postposition, each conjunction, etc., should be written as a different word. The grammar of the language, however, raises complications in the straightforward application of this principle.

One example of this is with possessives. Possession is indicated in Anii with a possessive pronoun between the possessor and the possessee, as in (1), where the possessive pronoun is written separately for clarity. All data in this paper is given in IPA, with accents indicating tone:

- (1) bàsm̀pìrì bá b̀ə̀fě̀i
 women POSS baskets
 The women's baskets

The community consensus was initially that it would be simpler to follow the semantics and write each word separately. However, the phonology of these constructions creates problems with this approach.

For example, when the thing being possessed begins with a vowel, the vowel of the possessive marker deletes, but it's tone remains. Additionally, in some cases there is /g/-deletion and compensatory lengthening, as well. Examples of these complications are shown step by step in (2). The phrase in (2a) means the women's shirts, that in (2b) your (sg) basket

- (2) a. Underlying Form: bàsm̀pìrì bá itókò
 women POSS shirts
 Possessive vowel deletion: bàsm̀pìrì b' itókò
 Tonal Reattachment: bàsm̀pìrì b itókò
 Surface Form: bàsm̀pìrì b itókò
 b. Underlying Form: àkì g̀ə̀fě̀i
 you.SG.POSS basket
 [g]-deletion: àkì àfě̀i
 Compensatory lengthening: àkì ààfě̀i
 Possessive vowel deletion: àk ààfě̀i
 Surface form: àkààfě̀i

The best way to write forms like these was the most hotly-contested issue at the orthography meeting. Due to the complex phonology of these forms, if the possession were to be written separated from the possessed noun, it is unclear where the space should be. The form in (2a) could be written b itoko, bi toko or bitoko, and (2b) could be written aka ashɛɪ, ak aashɛɪ, akaa shɛɪ, akaashɛɪ or even akɪgashɛɪ, though no speakers ever accepted the writing of the underlying form. It was data like this that finally led the meeting participants to decide that possessives should be written attached to the possessee noun, so that the phrase in (2a) is written 'basmpərə bitoko', and that in (2b) is written 'akaashɛɪ'. Thus, the linguistic facts of the language over-ruled the initial community consensus.

The harmony process means that the vowel of every noun class marker surfaces with the same [ATR] quality as the vowels of the noun root that follows it, i.e. there is neutralization of the [ATR] contrast in these positions. Noun-class agreement markers on verbs and adjectives also agree in [ATR] with the roots that follow them. This phenomenon is illustrated in (6) with every relevant noun class:

(6)	[-ATR]		[+ATR]	
a. Class A:	no known example		[ù-pì]	‘child’
b. Class Ə ¹ :	[à-bòrí]	‘sheep/animal’	[ə-kutu]	‘orange’
c. Class C:	[gà-filí]	‘fish (sg.)’	[gì-dú]	‘place’
d. Class Ð:	[gì-bó]	‘very short shorts’	[gì-dzè]	‘yam’
e. Class E:	[õ-dò]	‘neck’	[ù-ʃìnè]	‘heart/courage’
f. Class ɛ:	[gõ-tò]	‘ear’	[gù-jó]	‘tree’
g. Class G/ʊ ² :	[bõ-tòŋà]	‘salt’	[bù-tó]	‘water’
h. Class T ³ :	[à-né]	‘hands’	[ə-kú]	‘rooms/huts’
i. Class U:	[ì-lò]	‘strings/cords’	[ì-dzè]	‘yams’
j. Class W:	[ì-tòŋà]	‘guinea fowl (pl)’	[ì-bú]	‘snakes’
k. Class Y:	[bà-fòmì]	‘farmers’	[bè-pì]	‘children’

Unlike the nasal harmony discussed above, this vowel harmony process is not at all salient to speakers, and it turned out to be quite difficult to teach people to write it. At the orthography meeting, everyone thought it was a good idea to consistently write the vowels of noun-class prefixes or agreement markers the same way because most Anii-speakers are unable to accurately determine which vowel is present in these circumstances. Given that there is no linguistic evidence as to whether the [+ATR] or the [-ATR] forms are the underlying ones, the decision was taken to write the [-ATR] forms in all situations, based upon impressionistic ideas that these were the more common vowels in the language.

For years, there was confusion among writers as to which vowel to write in these situations. With the linguistic analysis to show what was happening, the confusion was easily explained and a non-confusing solution to this problem that everyone could agree on was approved.

3.0 Social factors

Some orthography decisions were based on social factors. That is, rather than using the form of the language to make orthography rules, the social milieu in which the language is spoken was taken into account.

3.1 Dialect choice

Since there are currently not enough resources to support developing an orthography for each of the Anii dialect areas, so it was decided to focus on the dialect of Bassila, even though it is not linguistically the most central dialect for the following reasons:

- 1 The twelve Anii noun classes are currently designated with letters from the Anii alphabet, with the ‘singular’ classes given letters from the beginning of the alphabet and the ‘plural’ classes letters from the end (Zaske 2011). Class B has no noun-class prefix and Class F has no vowel (the noun-class marker is a nasal consonant), so they are not mentioned here.
- 2 These two classes have been kept apart in the past because Class G contains singular objects or mass nouns and Class ʊ plural objects, and the co-authors are still debating whether that separation should remain.
- 3 Classes Ə and T, and U and W respectively have the same noun-class markers but trigger different agreement patterns.

- Bassila is the largest Anii village (actually more like a town)
- The local radio station broadcasts in the Bassila dialect, and people in most other villages listen to that
- People from all the villages present at the orthography meeting agreed that adults in every village understand the Bassila dialect.

Representatives from all the villages agreeing on these points was a key factor in the decision-making process, since the speakers of the Bassila dialect were uncomfortable making this choice for speakers of other dialects. Linguistic analysis would have chosen a more northern dialect to be the written form of the language, since those are linguistically the most central dialects, but the social situation required a different choice.

3.2 The writing of the phoneme [d]

In the official Anii orthography, the phoneme [d] is written ‘d’. Linguistically-speaking, there is no voiced retroflex stop [ɖ] in the language, there is only the voiced alveolar stop [d]. However, for a number of historical and social reasons, the phoneme [d] continues to be written as ‘d’. These reasons are:

- The anthropologist who originally helped develop the alphabet in the 1970s thought that he heard the phoneme [ɖ], though this was likely a mis-hearing, given that the Anii speaker he was working with is still alive today (and active in language development), and we have not heard this pronunciation from him.
- The old gentlemen who have been involved in language development for over 40 years are used to writing the phoneme [d] as ‘d’ and find it aesthetically pleasing, and the language committee also has insisted on keeping this letter for many years.
- Since some neighboring languages have both [d] and [ɖ] as phonemes, while French, the lingua franca of the area, only has [d], some speakers argue that using ‘d’ in writing Anii makes it easier to learn to read surrounding local languages. This is not a linguistic argument, but rather a social one.
- Since there have been many publications since 2001 which use the d, and the SIL publication team has set up their keyboards and other equipment with that letter, it would be very labor-intensive to change that decision now.

As can be seen, these are largely social and historical arguments. The linguistic data would suggest that the phoneme [d] should be written as ‘d’, but the social arguments prevailed, and the the decision was taken that the phoneme should be written ‘d’.

4.0 The interaction of linguistic and social factors

While all orthography decisions involve both linguistic and social factors, some decisions cannot be argued to be largely one or the other, but show clear interactions. These are generally cases where the simplest solution linguistically-speaking did not work for social reasons.

4.1 The eleventh vowel and syllabic consonants

Bassila Anii has eleven vowel phonemes, as shown above, but the other dialects, as far as is known, have only ten, they do not have [i]. The existence of [i] as a separate phoneme from [ə] was only discovered in 2009 after extensive linguistic analysis (see Morton 2012), during which it was also

discovered that Gɪsɪda Anii has syllabic nasals and liquids. Before this linguistic analysis, there was confusion in reading caused by many different words being written the same way, as shown by the example in (7) below:

(7)	<i>transcription</i>	<i>gloss</i>	<i>original spelling</i>	<i>proposed spelling (after linguistic analysis)</i>
	[pə̀l]	go around	pəl	pəl
	[pɪl]	be cooked	pəl	pil
	[pɪ]	pursue	pəl	pl

The initial suggestion for solving this problem, based on the linguistic data, was to write the [ɪ] as 'i', as is done in one nearby language with a similar vowel system, and to not write a vowel where linguistically there is no vowel, as illustrated in (7) above. The argument was that this system could be easily taught to speakers of other dialects because there are regular vowel correspondances among dialect areas.

Those present at the orthography meeting initially liked this proposal, except that in monosyllabic words, they thought that the absence of a vowel looked too strange and would be hard to read. Additionally, they had a different take on the dialect issue, and decided that since there was less confusion in many non-Bassila dialects about this issue, they should base some decisions on the facts of those dialects. After much discussion, the following rules were agreed on:

- The phoneme [ə], which is found cross-dialectally, will be written 'ə'
- The phoneme [ɪ] will also generally be written as 'ə', as has been done historically, but where this created confusion, as in (5), the phoneme will be written as 'ɪ', which is how it is generally pronounced in Northern dialects.
- In words longer than one syllable, syllabic consonants will be written with no vowel (for example, [asɲa], 'dog', will be written 'asna'. With one-syllable words, though, the syllabic consonants will be written preceded by an 'i', as those words are pronounced in Northern dialects, i.e. [pɪ] will be written as 'pil'.

The details of the spelling of each word are still being worked out, but these general rules are working well though they are much more complicated than the original proposal, due largely to social considerations such as dialect variation. One note is that in monosyllabic vowel-final words, these rules may not be enough to disambiguate words which are currently homographs. An example is given at the end of the next section.

4.2 Tone

Anii has two phonemic tones, high (H) and low (L), with a phonetic mid tone caused by downstep in certain contexts. There is both grammatical and lexical tone. If no tone markings are written, confusion can occur, as illustrated in (8), with examples of grammatical tone given in (8a) and lexical tone in (8b):

(8)	<i>transcription</i>	<i>gloss</i>	<i>orthography without tone</i>
a.	[ɲ tɪ ! kpá]	I am arriving	n tɪ kpa
	[ɲ ! tɪ ! kpá]	I eventually arrived	n tɪ kpa
	[gɪ ɲɔ̀n]	We wrote	gɪ ɲɔ̀n
	[gɪ ! ɲɔ̀n]	Let's write	gɪ ɲɔ̀n
	[kɪ má kpá ná]	I did not arrive	kə ma kpa na
	[kɪ ! má kpá ná]	I will not arrive	kə ma kpa na

b. [m̀pà ǹ dà]	The palm nut is there	mpa n da
[m̀pá ǹ dà]	The village is there	mpa n da
[n̄ sàà]	I reserved	n saa
[n̄ sàá]	I waited	n saa
[à kpá]	He is about to arrive	a kpa
[á kpá]	You (sg) are about to arrive	a kpa

Because of examples like these, it was clear to everyone that tone needed to be written, at least in some cases, but there was a very strong community value against writing a lot of tone marks, because it made the texts harder to read (or at least led them to be perceived as being harder to read), and also was not aesthetically pleasing.

The final conclusion of the orthography meeting was that lexical tone should not be written because the context can generally disambiguate the words without tone marking, and the number of tone marks that should be written was limited. The major exception to this is the pronoun example given last in (8b), because there were too many contexts that did not disambiguate this pair. Grammatical tone, however, can be, but generally is not disambiguated by context, and therefore is written. Given these rules, the examples in (8) above are now written as follows:

(9)	<i>transcription</i>	<i>gloss</i>	<i>current orthography (with tone)</i>
a.	[n̄ tí ! kpá]	I am arriving	n tí kpa
	[n̄ ! tí ! kpá]	I eventually arrived	n tí kpa
	[gí ɲòn]	We wrote	gɪ ɲon
	[gí ! ɲón]	Let's write	gɪ ɲón
	[kí má kpá ná]	I did not arrive	kə ma kpa na
	[kí ! má kpá ná]	I will not arrive	kə ma kpá na
b.	[m̀pà ǹ dà]	The palm nut is there	mpa n da
	[m̀pá ǹ dà]	The village is there	mpa n da
	[n̄ sàà]	I reserved	n saa
	[n̄ sàá]	I waited	n saa
	[à kpá]	He is about to arrive	a kpa
	[á kpá]	You (sg) are about to arrive	á kpa

There remain several places where confusion remains, given the decisions above. Most pressing is how to write the difference between [tsí], 'go' and [tsə̀], 'build'. Neither the vowel rules given in the previous section, nor the tonal rules above can easily distinguish between these two words, both currently written 'cə̀', and context does not always distinguish easily either. As the orthography is used more and more, it is possible that specific cases like these will have to be addressed by amending the orthography rules as needed, particularly in the process of creating an Anii-French dictionary, which is currently in progress.

5.0 Conclusion and applications

The examples given above from the process of developing the Anii orthography have been given to illustrate two important points about orthography development. First of all, orthography decisions cannot be made without solid linguistic research (including, preferably, some knowledge of dialect differences) so that the orthography reflects (at least to some extent) the facts of the language, and is thus reasonably easy to learn and use. Secondly, however, such decisions also cannot be made on purely linguistic grounds. Orthography developers must also take into account the community's language

attitudes and meta-linguistic awareness, since these things also have significant effects on a community's ability and willingness to use an orthography.

The case of Anii suggests that orthography development for unwritten languages is best carried out (when possible) by a collaboration between trained linguists and interested (and influential) members of the language community.

Works Cited

- Morton, Deborah. 2012. "[ATR] Harmony in an Eleven-Vowel Language: The Case of Anii". In: edited Marlo, Michael R., Nikki B. Adams, Christopher R. Green, Michelle Morrison, and Tristan M. Purvis, Eds. *Selected Proceedings of the 42nd Annual Conference on African Linguistics: African Languages in Context*. Somerville, MA: Cascadilla. On-line at: <http://www.lingref.com/cpp/acal/42/abstract2759.html>
- Zaske, Stefanie and Hakimou ATTI KALAM. 2011. *Écriture de la Langue Anii, Tome 3: Proposition d'une Orthographe Préliminaire pour la Langue Anii*. Bassila: SIL Bassila.

Note: A four-volume series (written in French) on the Anii orthography, including history, linguistic examples, and the final orthography rules, is available from the second author upon request.

Acknowledgements

Many thanks to all the Anii community members who gave of their time and their wisdom to help with the development of the orthography, especially those present at the orthography meeting, the members of the research group who continue to refine and apply the orthography, and the SIL team. Thanks in particular go to ATTI KALAM Hakimou, who has worked for many years on the linguistic research that made the Anii orthography possible. Special recognition also goes to the 'Fathers of Written Anii', IBRAHIMA GOMINA Mahazou, KALAM Karim, and GOMON Idikoukou Yacoubou (pictured below at a research group meeting in 2012), who worked on the initial development of the Anii alphabet and orthography, and are still involved in the development of their language.

